



Analisis Sentimen Review Film Avatar 2 pada Platform IMDb Menggunakan LSTM dan GRU

Rani Saputri^{1*}, Anna Baita²

¹⁻² Universitas Amikom Yogyakarta, Indonesia

Jl.Padjajaran, Ring Road Utara, Kel.Condongcatur, Kec.Depok,
Kab.Sleman, Prop.Daerah Istimewa Yogyakarta 55283

E-mail: ranisaputri@students.amikom.ac.id¹, anna@amikom.ac.id²

Abstract. This research uses a deep learning-based sentiment analysis approach with several main stages, namely data collection, preprocessing, model preparation, and model building. In addition, this research also evaluates the impact of data splitting techniques on the model's performance during the training process. The evaluation results show that the LSTM-GRU model achieved the best performance on the character aspect, with an F1-score of 0.72 in the 90:10 splitting scenario. Meanwhile, the lowest F1-score was found in the editing aspect, with a value of 0.51 in the 80:20 splitting scenario. These findings indicate that the model is more effective in recognizing sentiment in narrative aspects compared to technical aspects. Furthermore, the data splitting technique significantly influences model performance, both in determining accuracy levels and in optimizing the model's effectiveness in identifying sentiment patterns more accurately.

Keywords: Sentiment Analysis, LSTM, GRU, Imbalance Data, Movie Reviews

Abstrak. Penelitian ini menggunakan pendekatan *sentiment analysis* berbasis *deep learning* dengan beberapa tahapan utama, yaitu pengumpulan data (*data collection*), pra-pemrosesan (*preprocessing*), persiapan model (*model preparation*), dan pelatihan model (*model building*). Selain itu, penelitian ini mengevaluasi pengaruh teknik *data splitting* terhadap performa model selama proses pelatihan. Hasil evaluasi menunjukkan bahwa model LSTM-GRU mencapai performa terbaik pada aspek *character* dengan F1-score sebesar 0.72 dalam skenario *splitting* 90:10, sedangkan F1-Score terendah untuk pengujian model LSTM dan GRU ada pada aspek editing dalam skenario 80:20 dengan nilai 0.51.. Hal ini menunjukkan bahwa model lebih efektif dalam mengenali sentimen pada aspek naratif dibandingkan aspek teknis. Selain itu, teknik *data splitting* terbukti berpengaruh terhadap performa model, baik dalam menentukan tingkat akurasi maupun efektivitas model dalam mengenali pola sentimen secara lebih optimal.

Kata kunci: Analisis Sentimen, LSTM, GRU, Imbalance Data, Ulasan Film

1. LATAR BELAKANG

Film adalah salah satu bentuk hiburan yang sangat digemari di seluruh dunia (Ramadhan et al., 2024). Film juga dapat menjadi media yang dapat menyampaikan pesan moral, sosial dan emosional kepada audiens (Sagita et al., 2024). Oleh karena itu, analisis terhadap ulasan film sangat penting untuk mengetahui bagaimana reaksi penonton dalam memandang kualitas sebuah film dari berbagai aspek. Internet Movie Database (IMDb) adalah platform terkemuka yang menyediakan informasi lengkap tentang film dan memungkinkan pengguna memberikan rating serta ulasan (Cahyani et al., 2022).

Film avatar 2 yang dirilis secara global dan mendapatkan rating tertinggi ke-4 sepanjang masa di tahun 2023 (*Box Office: "Avatar 2" Kalahkan "Last Jedi" Di Daftar Box Office AS Sepanjang Masa*, n.d.). Film Avatar 2 (*The Way of Water*) karya James Cameron, dirilis pada 14 Desember 2022 (Pandunata et al., 2023). Penilaian dari IMDb serta gambaran tanggapan penonton terkait dengan aspek *character, mise en scene, time, location, conflict*,

Received: Februari 05, 2025; Revised: Februari 20, 2025; Accepted: Maret 03, 2025; Published : Maret 06, 2025

cinematography, editing, sound dapat dilakukan dengan *sentiment analysis*(Pandunata et al., 2023). *Deep Learning* menjadi topik populer dalam *sentiment analysis* dengan arsitektur yang sering digunakan meliputi *Convolutional Neural Network* (CNN)(Adam & Setiawan, 2023), *Long Short Term Memory*(LSTM)(Nurvania & Lhaksamana, 2021), dan *Recurrent Neural Network* (RNN)(Utami, 2022).

CNN tidak hanya dapat memproses gambar tetapi juga dapat memproses teks, tetapi sulit menangani dependensi jangka panjang(Adam & Setiawan, 2023). RNN memungkinkan pemrosesan teks menjadi lebih efektif dengan memanfaatkan informasi sebelumnya untuk menghasilkan analisis sentimen yang lebih akurat(Merinda Lestandy et al., 2021). Namun RNN, memiliki keterbatasan dalam menangani teks yang terlalu panjang(Jabat et al., 2024). Untuk mengatasi kekurangan ini dibentuklah *Long Short Memory* (LSTM), jaringan LSTM mampu menerima input dari waktu saat ini dan output dari waktu sebelumnya(Mohammadi & Shaverizade, 2021). Untuk itu penelitian ini menggunakan metode LSTM yang dikombinasikan GRU.

Selain arsitektur *deep learning*, metode *splitting* juga dapat mempengaruhi performa model. Seperti penelitian yang dilakukan oleh Riskawati(Riskawati et al., 2024). Kemudian proses ekstraksi fitur dalam *deep learning* dapat dilakukan dengan *word embedding*. Seperti penelitian yang dilakukan oleh Poetra(Poetra et al., 2022)dengan menggunakan *word embedding* dapat meningkatkan hasil akurasi model.

Dataset review film *Avatar The Way of Water* yang diperoleh melalui *scraping* data melalui platform IMDb memiliki karakteristik fitur dengan jumlah token yang besar. oleh karena itu penelitian ini menggunakan arsitektur turunan *Recurrent Neural Network* (RNN) dengan menggunakan Glove sebagai ekstraksi kata menjadi vektor, serta dilakukan pula uji *splitting*, untuk kualitas model yang baik. Penelitian ini bertujuan untuk mengukur seberapa baik kinerja model *Long Short Term Memory* (LSTM) yang dikombinasikan dengan *Gated Recurrent Network* (GRU) dalam melakukan analisis berbasis aspek terhadap ulasan pengguna platform IMDb berdasarkan setiap aspek untuk film “*Avatar The Way of Water*”

2. KAJIAN TEORITIS

IMDb

IMDb, atau *Internet Movie Database*, adalah platform yang menyediakan informasi lengkap tentang film, acara TV(Tobing, 2024), serta aktor yang terlibat, sekaligus mengumpulkan ulasan, penilaian, dan opini dari pengguna yang ingin mengetahui lebih banyak berdasarkan pengalaman penonton lainnya(Ramadhan et al., 2024). IMDb berperan penting

dalam industri perfilman sebagai sumber informasi utama yang memengaruhi keputusan dan minat penonton. Platform ini membantu calon penonton menilai kualitas sebuah film sebelum mereka memutuskan untuk menontonnya(Aziz Bawazir & Daniel, 2024).

Sentiment Analysis

Analisis sentimen, atau yang dikenal juga dengan penambahan opini, adalah studi yang bertujuan menganalisis opini masyarakat terhadap produk, layanan, individu, atau topik tertentu(Hadi, n.d.). Proses ini dilakukan secara otomatis untuk memahami sentimen dalam data teks tidak terstruktur, yang dapat digunakan di berbagai bidang seperti ekonomi, politik, sosial, dan hukum(Arsi & Waluyo, 2021).

Preprocessing

Teknik preprocessing merupakan serangkaian langkah yang penting untuk mempersiapkan data mentah agar siap dianalisis, terutama dalam hal persiapan data untuk digunakan dalam suatu proyek atau sistem(Fan et al., 2021). Teknik preprocessing mencakup beberapa langkah utama, seperti *Clean Text*, *Case Folding*, *Lemmatization*, *Stopword Removal*, *Lemmatization*, EDA(Amelia & Aminuallah, 2023).

Augmentasi

Augmentasi data menambah variasi dalam dataset untuk meningkatkan jumlah sampel pelatihan. Teknik ini membantu model belajar dari lebih banyak contoh, mengurangi overfitting, dan meningkatkan generalisasi terhadap data baru. Dalam analisis sentimen dan pembelajaran mesin, augmentasi data membuat model lebih robust dan akurat.

Word Embedding Glove

Word embeddings adalah proses mengubah teks menjadi representasi numerik agar dapat dipahami oleh algoritma pembelajaran mesin dan deep learning(Riyantoko et al., 2022). Salah satu metode yang digunakan adalah GloVe, yang mengonversi kata ke dalam bentuk vektor dengan pendekatan *global matrix factorization*, yaitu pemfaktoran matriks yang merepresentasikan kemunculan atau ketiadaan kata dalam suatu dokumen. Representasi vektor ini sangat penting dalam analisis sentimen, karena memungkinkan data yang dikumpulkan diolah secara lebih akurat sebelum dianalisis lebih lanjut(Setiawan & Setiawan, 2023).

Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN) adalah algoritma jaringan saraf buatan yang dirancang khusus untuk mengolah data dalam bentuk urutan atau sekuens(Gultom & Siregar, 2024). RNN dirancang untuk mengenali pola dalam data berurutan dengan menyimpan informasi dari langkah sebelumnya melalui mekanisme memori. Berbeda dengan jaringan saraf feedforward, RNN efektif dalam pemrosesan bahasa alami, analisis deret waktu, dan

pengenalan suara.

Long Short Term Memory (LSTM)

Long Short-term Memory (LSTM) adalah arsitektur jaringan saraf tiruan yang dikembangkan dari *Recurrent Neural Network* (RNN) untuk mengatasi masalah hilangnya efektivitas gradien (Hadi, n.d.). Metode LSTM ini sangat cocok dan efektif digunakan dalam proses klasifikasi teks, karena mampu menghasilkan performa yang baik pada teks berbahasa Indonesia maupun Inggris (Poetra et al., 2022).

Long Short-Term Memory (LSTM) dirancang untuk mengatasi keterbatasan *Recurrent Neural Network* (RNN) dalam menangani urutan panjang dengan menggunakan sel memori yang menyimpan informasi lebih lama, sehingga dapat menangkap dependensi data yang melibatkan banyak langkah waktu (Yudi Widhiyasa et al., 2021). Dengan mekanisme ini, LSTM mengatasi masalah *vanishing gradient* yang sering muncul pada RNN tradisional, menjadikannya lebih efisien dalam memproses data sekuensial seperti teks dan suara.

Gated Recurrent Unit (GRU)

GRU adalah varian LSTM dengan desain lebih sederhana dan perhitungan lebih ringan, tetapi tetap akurat. Model ini memiliki dua gerbang utama: *update gate* untuk menyaring informasi yang dipertahankan dan *reset gate* untuk mengabaikan informasi lama. Sistem ini membantu mengatasi hilangnya *gradien*, menjadikannya efisien dalam menangani data berurutan (Adam & Setiawan, 2023).

Evaluasi

Evaluasi merupakan proses penilaian yang bertujuan untuk mengukur kualitas, kinerja, serta efektivitas suatu hal. Dalam melakukan evaluasi klasifikasi data, beberapa parameter yang umum digunakan meliputi *accuracy*, *precision*, *recall*, dan *F1-Score* (Riyantoko et al., 2022). Parameter evaluasi berperan dalam menilai efektivitas sistem, mengukur keberhasilannya, dan mendukung pengambilan keputusan untuk perbaikan serta optimalisasi. Evaluasi juga membantu strategi pengembangan guna meningkatkan kualitas sistem.

Accuracy: Seberapa sering model membuat prediksi yang benar dibandingkan dengan seluruh data yang diuji.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

precision : Mengukur seberapa banyak prediksi positif yang benar dibandingkan dengan semua yang diprediksi positif.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

recall: Mengukur seberapa banyak data positif yang benar-benar terdeteksi dibandingkan dengan semua data positif yang ada.

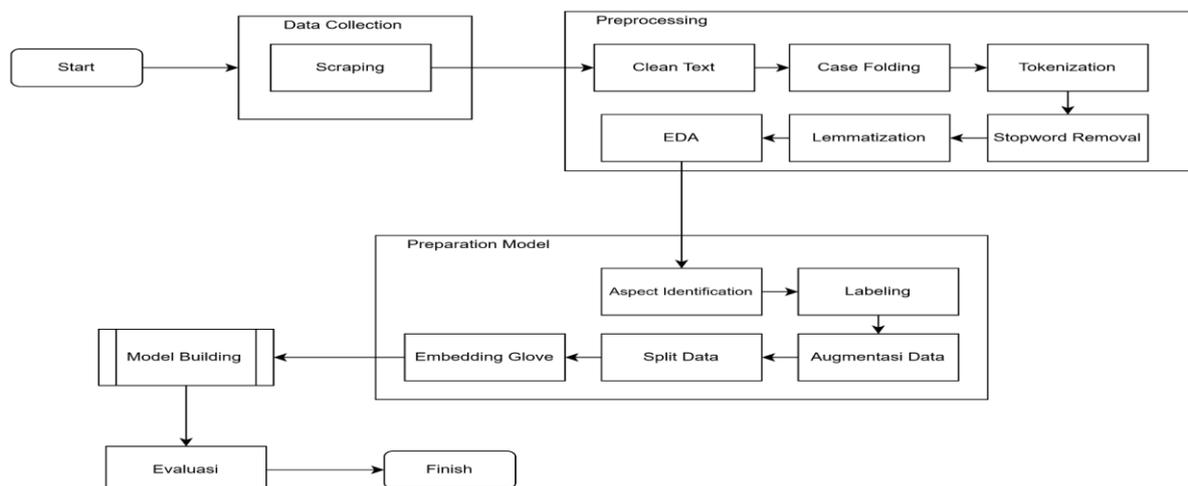
$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

F1-score : rata-rata harmonik dari presisi dan recall, digunakan untuk keseimbangan antara keduanya.

$$\text{F1-Score} = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

3. METODE PENELITIAN

Penelitian ini dirancang dengan serangkaian langkah percobaan yang sistematis guna mencapai target penelitian yang telah ditetapkan. Alur Penelitian dapat di lihat di gambar 1 berikut.



Gambar 1 Alur Penelitian

- **Data Collection:** Data ulasan untuk film *Avatar The Way of Water* berhasil dikumpulkan dari situs IMDb menggunakan teknik *web scraping*. Sebanyak 3.223 ulasan berhasil diperoleh, yang selanjutnya akan diproses dan disimpan untuk tahap *preprocessing*.
- **Preprocessing:** Teknik preprocessing merupakan serangkaian langkah yang penting untuk mempersiapkan data mentah agar siap dianalisis, terutama dalam hal persiapan data untuk

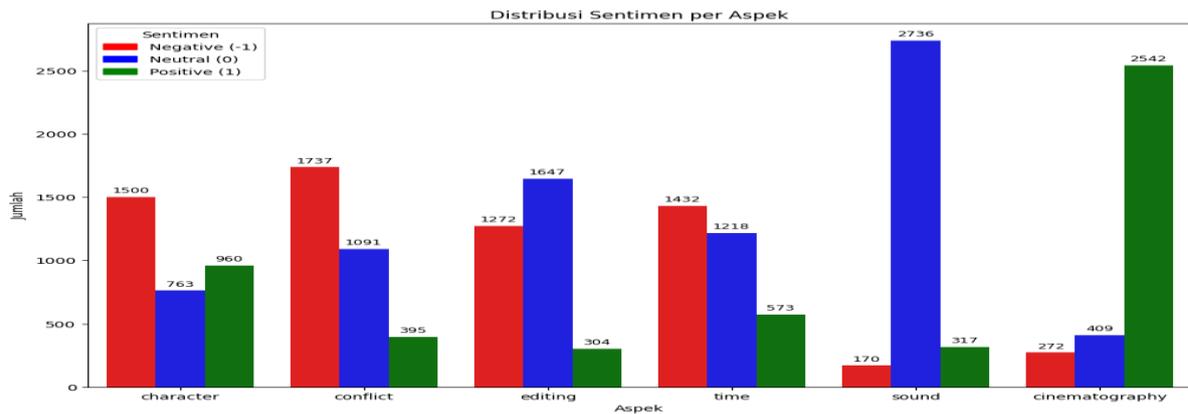
digunakan dalam suatu proyek atau sistem (Fan et al., 2021). Teknik preprocessing mencakup beberapa langkah utama, seperti *Clean Text*, *Case Folding*, *Lemmatization*, *Stopword Removal*, *Lemmatization*, EDA (Amelia & Aminullah, 2023).

- *Clean Text* untuk membersihkan teks dengan menghapus beberapa elemen tertentu yang biasanya terdapat dalam sebuah ulasan.
 - *Case Folding* untuk menormalisasikan teks dan menyamakan format data.
 - *Tokenization* memecah teks menjadi kata-kata atau token.
 - *Stopword Removal* menghapus kata-kata yang terdapat di *stoplist* yang berisi kata-kata konjungsi.
 - *Lemmatization* mengubah kata-kata ke dalam bentuk dasarnya.
 - EDA
- **Preparation Model:** Dalam persiapan model, terdapat tiga proses utama, yaitu augmentasi, pemisahan data, dan embedding Glove. Augmentasi dilakukan menggunakan TextBlob untuk memperkaya teks melalui sinonimisasi, perubahan struktur kalimat, dan modifikasi kata berdasarkan konteks tanpa mengubah makna aslinya. Pemisahan data diterapkan dalam empat skenario rasio, yaitu 60:40, 70:30, 80:20, dan 90:10, guna mengevaluasi pengaruh jumlah data latih terhadap akurasi dan generalisasi model. Embedding Glove menggunakan varian word embedding Gigaword dengan dimensi vektor 200 untuk merepresentasikan kata dalam bentuk vektor yang lebih mendalam, meningkatkan pemahaman model terhadap makna kata. Dalam penelitian ini, aspek yang digunakan terdiri dari enam kategori, yaitu character, conflict, editing, time, sound, dan cinematography. Data kemudian diklasifikasikan secara manual berdasarkan sentimen positif (1), negatif (-1), dan netral (0), serta dianalisis berdasarkan aspek-aspek tersebut.
 - **Model Building:** Model yang dibangun menggunakan *arsitektur Long Short-Term Memory* (LSTM) yang dikombinasikan dengan *Gated Recurrent Unit* (GRU), dengan fungsi aktivasi sigmoid dan optimasi menggunakan algoritma Adam. Untuk mengukur tingkat kesalahan selama pelatihan, model ini menerapkan fungsi *loss sparse_categorical_crossentropy* agar dapat melakukan klasifikasi lebih efisien.
 - **Evaluasi:** Evaluasi dilakukan menggunakan metrik seperti *accuracy*, *precision*, *recall*, dan *f1-score* untuk menganalisis performa model secara lebih mendalam dan memastikan efektivitasnya dalam klasifikasi.

4. HASIL DAN PEMBAHASAN

Profil Dataset

Penelitian ini menggunakan 3.223 data yang telah divisualisasikan dalam gambar yang dapat dilihat di lihat di gambar 2. Data tersebut sebelumnya telah melalui proses pelabelan manual dan diklasifikasikan ke dalam enam aspek, yaitu *Character*, *Conflict*, *Editing*, *Time*, *Sound*, *Cinematography*.



Gambar 2 Barplot

Dapat dilihat pada Gambar 2 bahwa data yang dimiliki tidak seimbang. Sentimen negatif terbanyak terdapat pada aspek *conflict* dengan jumlah tertinggi (1737), diikuti oleh aspek *character* (1500) dan aspek *time* (1432). Selain itu, aspek *sound* lebih dominan dalam sentimen netral (2736). Sementara itu, aspek *cinematography* lebih banyak mendapatkan sentimen positif (2542), disusul oleh aspek *character* (960) dan *time* (573). Ketidakeimbangan ini menunjukkan bahwa setiap aspek dalam dataset memiliki distribusi sentimen yang berbeda, dengan beberapa aspek lebih cenderung mendapat sentimen tertentu dibandingkan yang lain. Berikut tabel 1 komposisi data per aspek.

Skenario Pengujian

Berikut skenario pengujian yang dapat dilihat di tabel 1.

Tabel 1 Skenario pengujian

Unit neuron	256
Dropout	0.4 dan 0.5
Dense	128
Aktivasi	Relu dan Sigmoid
Loss	Sparse_categorical_crossentropy
Kernel_Regularizer	L2(0.03)

Optimizer	Adam
Learning Rate	0.00001
Epoch	100
Early Stopping	5

Berdasarkan Tabel 1, semua skenario pengujian memakai komponen yang sama. Setiap model menggunakan 256 unit neuron, menerapkan dropout sebesar 0.4 dan 0.5 untuk mencegah overfitting, serta memiliki lapisan Dense dengan 128 unit neuron menggunakan fungsi aktivasi ReLU dan Sigmoid. Fungsi loss yang digunakan adalah *sparse_categorical_crossentropy*, cocok untuk klasifikasi multi-kelas dan multi-label. Model menerapkan regulasi L2 dengan nilai 0.3 untuk mencegah kompleksitas berlebih. Optimasi dilakukan dengan Adam menggunakan learning rate 0.00001, dilatih selama 100 epoch dengan *early stopping* sebanyak 5 epoch jika tidak ada peningkatan performa.

Hasil Perbandingan

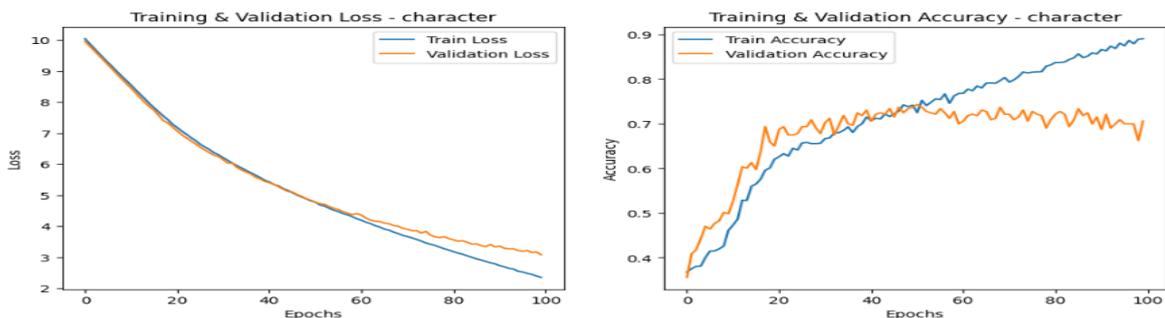
Hasil perbandingan Tabel 2 menggunakan nilai F1-Score sebagai acuan untuk menentukan model terbaik diantara empat skenario pengujian.

Tabel 2 Hasil Perbandingan Pengujian Model

Aspek	Skenario	SVM	LSTM	LSTM + GRU	LSTM Multi Aspek
Character	60:40	0.56	0.53	0.67	0.60
	70:30	0.57	0.52	0.64	0.60
	80:20	0.58	0.62	0.70	0.56
	90:10	0.61	0.66	0.72	0.62
Conflict	60:40	0.46	0.47	0.57	0.51
	70:30	0.46	0.53	0.56	0.58
	80:20	0.56	0.44	0.64	0.56
	90:10	0.48	0.56	0.62	0.50
Editing	60:40	0.45	0.36	0.52	0.47
	70:30	0.47	0.37	0.55	0.47
	80:20	0.48	0.41	0.51	0.51
	90:10	0.45	0.41	0.56	0.51

Aspek	Skenario	SVM	LSTM	LSTM + GRU	LSTM Multi Aspek
Time	60:40	0.52	0.42	0.55	0.52
	70:30	0.53	0.42	0.54	0.54
	80:20	0.49	0.44	0.57	0.55
	90:10	0.50	0.51	0.58	0.54
Sound	60:40	0.51	0.35	0.52	0.45
	70:30	0.53	0.28	0.56	0.39
	80:20	0.55	0.45	0.58	0.44
	90:10	0.56	0.44	0.57	0.42
Cinematography	60:40	0.49	0.36	0.52	0.39
	70:30	0.50	0.31	0.52	0.36
	80:20	0.49	0.44	0.53	0.29
	90:10	0.48	0.59	0.52	0.36

Tabel 2 Dari hasil perbandingan empat model, dapat dilihat bahwa LSTM + GRU menghasilkan F1-score tertinggi di hampir semua aspek, skenario dengan F1-Score tertinggi ada di aspek *character* dengan hasil 0.72, ini menunjukkan bahwa kombinasi LSTM dan GRU lebih efektif dalam menangkap pola sentimen. Dapat dilihat grafik performa pada gambar 3 tersebut.



Gambar 3 Grafik Aspek *Character* LSTM dan GRU

Gambar 3 menunjukkan bahwa model masih mengalami overfitting, terlihat dari validation accuracy yang cenderung menurun setelah epoch 50 sementara training accuracy terus meningkat. Namun, kelebihan yang dapat ditonjolkan adalah model berhasil mencapai akurasi validasi tertinggi sekitar 0.7 dan mempertahankannya dengan fluktuasi kecil. Selain itu, loss untuk training dan validasi menunjukkan tren penurunan yang stabil, menandakan

bahwa model mampu belajar dari data dengan baik meskipun masih perlu peningkatan dalam generalisasi.

Sementara itu model SVM dengan memakai skenario dari penelitian pandunata(Pandunata et al., 2023) menunjukkan performa yang cukup stabil di berbagai skenario, bahkan dalam beberapa aspek seperti *Character* dan *Conflict*, dengan didapatkan skenario tertinggi F1-Scorenya ada di aspek *character* dengan nilai 0.61, ini menandakan performanya lebih baik dibandingkan LSTM dan LSTM Multi-Aspek.

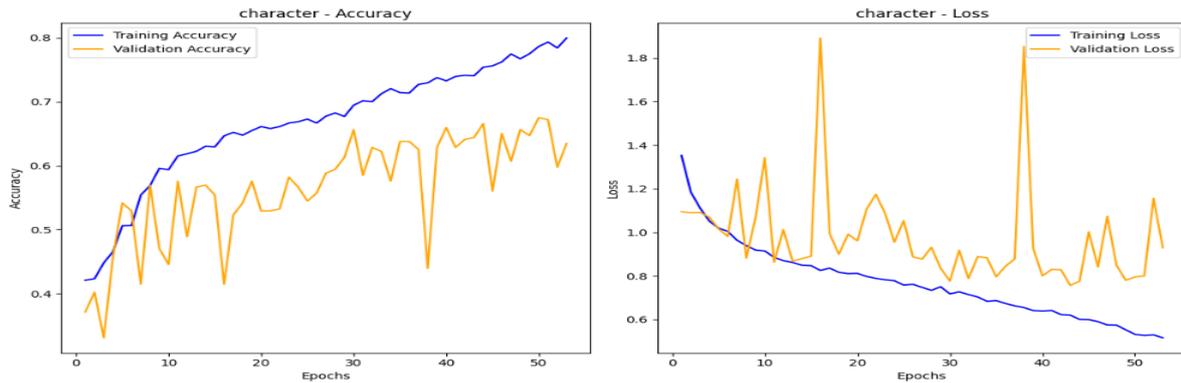
LSTM standar memiliki performa yang bervariasi, dengan F1- Score tertinggi ada di aspek *character* dengan nilai 0.66, namun, beberapa skenario menunjukkan skor yang lebih rendah dibanding SVM terutama pada aspek *Sound* dan *Cinematography*. Dapat dilihat dari grafik performa pada gambar 4 berikut.



Gambar 4 Grafik Aspek *Character*

Gambar 4 menunjukkan bahwa model masih mengalami overfitting, terlihat dari gap antara train accuracy dan validation accuracy setelah epoch 50. Namun, kelebihan yang dapat ditonjolkan adalah validation accuracy tetap mengalami peningkatan dan mampu mencapai sekitar 0.7 dengan fluktuasi yang lebih stabil dibandingkan grafik sebelumnya. Selain itu, loss untuk training dan validasi menurun dengan pola yang hampir sejajar, menandakan bahwa model masih mampu belajar dengan baik.

Sementara itu, LSTM Multi-Aspek tidak selalu unggul dibandingkan model lainnya. F1-Score tertinggi diperoleh aspek *character* dengan nilai 0.62, namun demikian, dalam beberapa Skenario, skornya lebih rendah dibandingkan model individual, menunjukkan bahwa mempertimbangkan semua aspek secara bersamaan belum tentu meningkatkan performa klasifikasi. Dapat disimpulkan bahwa model LSTM dengan GRU lebih baik di antara ketiga model diatas. Grafik performa dapat di lihat di gambar 5.



Gambar 5 Grafik Aspek *Character* Pada Multi Aspek

Gambar grafik 5 dengan skenario 90:10, masih terjadi *overfitting* yang cukup parah, terlihat dari *training accuracy* yang terus meningkat tetapi *validation accuracy* mengalami fluktuasi yang besar dan tidak menunjukkan peningkatan yang stabil.

5. KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa penggunaan model *Long Short Term Memory* (LSTM) yang dikombinasikan dengan *Gated Recurrent Unit* (GRU) cukup baik dalam menganalisis sentimen ulasan film *Avatar The Way of Water*.

Melalui berbagai skenario pengujian model dan teknik *splitting* data, diperoleh F1-Score tertinggi pada pengujian model LSTM yang dikombinasikan dengan GRU dalam skenario *splitting* 90:10, dengan nilai 0.72 pada aspek *character*. Sedangkan, F1-Score terendah untuk pengujian model LSTM dan GRU ada pada aspek *editing* dalam skenario 80:20 dengan nilai 0.51.

Selain itu, teknik *splitting* data yang digunakan turut berpengaruh terhadap performa model selama proses pelatihan, baik dalam menentukan tingkat akurasi maupun dalam meningkatkan efektivitas model dalam mengenali pola sentimen secara lebih optimal.

Saran

- Mempertimbangkan metode pra-pemrosesan teks yang lebih efektif untuk meningkatkan akurasi model, terutama pada aspek dengan performa rendah seperti editing.
- Mengeksplorasi teknik hyperparameter tuning serta penggunaan arsitektur model yang lebih canggih, seperti BERT atau GPT, untuk meningkatkan pemahaman model terhadap konteks sentimen.
- Meneliti variasi dalam teknik data *splitting* untuk menganalisis dampaknya terhadap kestabilan dan kemampuan model dalam mengenali pola sentimen di berbagai skenario.

- Mengembangkan model hingga tahap deployment untuk menguji performanya dalam lingkungan nyata.

DAFTAR REFERENSI

- Adam, A. Z. R., & Setiawan, E. B. (2023). Social media sentiment analysis using convolutional neural network (CNN) and gated recurrent unit (GRU). *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, 9(1), 119–131. <https://doi.org/10.26555/jiteki.v9i1.25813>
- Amelia, D. S., & Aminuallah, N. C. (2023). Teks dan analisis sentimen pada chat grup Whatsapp menggunakan long short-term memory (LSTM). *Jurnal Informatika dan Ilmu Data*, 3.
- Arsi, P., & Waluyo, R. (2021). Analisis sentimen wacana pemindahan ibu kota Indonesia menggunakan algoritma support vector machine (SVM). *Jurnal Teknologi Informasi dan Ilmu Komputer*, 8(1), 147. <https://doi.org/10.25126/jtiik.0813944>
- Aziz Bawazir, L. A., & Daniel, M. F. (2024). The influence of IMDb ratings for *KKN Film in Dancing Village* on students' watching interests (Students of STIKOM Interstudi Jakarta). *Journal of Humanities Social Sciences and Business (JHSSB)*, 3(2), 431–440. <https://doi.org/10.55047/jhssb.v3i2.960>
- Box Office: “Avatar 2” kalahkan “Last Jedi” di daftar box office AS sepanjang masa. (n.d.). Retrieved December 18, 2024, from <https://deadline.com/2023/01/box-office-avatar-the-way-of-water-infinity-pool-cronenberg-1235243380/>
- Cahyani, G., Widayani, W., Anggita, S. D., Pristyanto, Y., Ikmah, I., & Sidauruk, A. (2022). Klasifikasi data review IMDb berdasarkan analisis sentimen menggunakan algoritma support vector machine. *Jurnal Media Informatika Budidarma*, 6(3), 1418. <https://doi.org/10.30865/mib.v6i3.4023>
- Fan, C., Chen, M., Wang, X., Wang, J., & Huang, B. (2021). A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. *Frontiers in Energy Research*, 9, 652801. <https://doi.org/10.3389/fenrg.2021.652801>
- Gultom, F., & Siregar, R. (2024). Implementasi recurrent neural network sebagai IDS terhadap serangan jaringan. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 12(2).
- Hadi, S. F. S. (n.d.). Analisis sentimen menggunakan recurrent neural network terkait isu Anies Baswedan sebagai calon presiden 2024. *Jurnal Teknologi dan Ilmu Komputer*, 10(2), 1682.
- Jabat, D. E. B., Sipayung, L. Y., Raih, K., & Dakhi, S. (2024). Penerapan algoritma recurrent neural networks (RNN) untuk klasifikasi *Ulos Batak Toba*. *Jurnal Informatika dan Sains Data*, 1(2).
- Lestandy, M., Abdurrahim, A., & Syafa'ah, L. (2021). Analisis sentimen tweet vaksin COVID-19 menggunakan recurrent neural network dan naïve Bayes. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(4), 802–808. <https://doi.org/10.29207/resti.v5i4.3308>

- Mohammadi, A., & Shaverizade, A. (2021). Ensemble deep learning for aspect-based sentiment analysis. *International Journal of Nonlinear Analysis and Applications*, 12(Special Issue). <https://doi.org/10.22075/ijnaa.2021.4769>
- Nurvania, J., & Lhaksamana, K. M. (2021). Analisis sentimen pada ulasan di TripAdvisor menggunakan metode long short-term memory (LSTM). *Jurnal Ilmiah Teknologi dan Komputer*, 8(4), 4124.
- Pandunata, P., Nurdiansyah, Y., & Alfina, F. D. (2023). Aspect-based sentiment analysis of *Avatar 2* movie reviews on IMDb using support vector machine. *E3S Web of Conferences*, 448, 02041. <https://doi.org/10.1051/e3sconf/202344802041>
- Poetra, C. K., Pane, S. F., & Fatonah, N. S. (2022). Meningkatkan akurasi long short-term memory (LSTM) pada analisis sentimen vaksin COVID-19 di Twitter dengan Glove. *Jurnal Telematika*, 16(2), 85–90. <https://doi.org/10.61769/telematika.v16i2.400>
- Ramadhan, N. J., Putri, V. A. P., & Riyadi, D. A. (2024). Eksplorasi analisis sentimen pada rating film IMDb: Pendekatan perbandingan menggunakan decision tree dan naive Bayes. *Jurnal Informatika dan Teknologi Komputer*, 4(3), 7273–7286.
- Riskawati, R., Fatihanursari, F., Iin, I., & Rizki Rinaldi, A. (2024). Penerapan metode Naïve Bayes classifier pada analisis sentimen aplikasi Gopay. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 346–353. <https://doi.org/10.36040/jati.v8i1.8699>
- Riyantoko, P. A., Fahrudin, T. M., Prasetya, D. A., Trimono, T., & Timur, T. D. (2022). Analisis sentimen sederhana menggunakan algoritma LSTM dan BERT untuk klasifikasi data spam dan non-spam. *Prosiding Seminar Nasional Sains Data*, 2(1), 103–111. <https://doi.org/10.33005/senada.v2i1.53>
- Sagita, D. I., Arthansa, R. M., & Sari, A. P. (2024). Komparasi analisis sentimen ulasan film *Avengers: Endgame* di IMDb menggunakan metode Naïve Bayes dan SVM. *STORAGE: Jurnal Ilmiah Teknik dan Ilmu Komputer*, 3(3), 156–166. <https://doi.org/10.55123/storage.v3i3.4026>
- Setiawan, R. T., & Setiawan, E. B. (2023). Sentiment analysis of BBCA stock price on Twitter data using LSTM and genetic algorithm optimization. *Sinkron*, 8(4), 2479–2489. <https://doi.org/10.33395/sinkron.v8i4.12825>
- Tobing, F. A. T. (2024). Sentiment analysis of IMDb movie reviews using recurrent neural network algorithm. *Jurnal Teknologi dan Ilmu Komputer*, 16(1).
- Utami, H. (2022). Analisis sentimen dari aplikasi Shopee Indonesia menggunakan metode recurrent neural network. *Indonesian Journal of Applied Statistics*, 5(1), 31. <https://doi.org/10.13057/ijas.v5i1.56825>
- Widhiyana, Y., Semiawan, T., Mudzakir, I. G. A., & Noor, M. R. (2021). Penerapan convolutional long short-term memory untuk klasifikasi teks berita bahasa Indonesia. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 10(4), 354–361. <https://doi.org/10.22146/jnteti.v10i4.2438>